Comparing Methods for Multivariate Classifier Training Variable Selection using the Search for the Scalar Top Quark as a Case Study. DENNIS MACKIN, Rice University, DØ COLLABORATION — We look for a way to automate the process of training variable selection when applying multivariate event classifiers to the search for new phenomenon in high energy physics experiments. The DØ collaboration recently completed a search for the Supersymmetric partner of the top quark in the two muons, two jets, and missing transverse energy final state. We use the Monte Carlo events representing the signal and the background from this search as the basis for our case study. We begin with the computationally expensive, $\mathcal{O}(2^n)$, method of testing the classifier for all variable combinations and then selecting the one combination which gives the best expected signal sensitivity. We then compare this "best" sensitivity to the sensitivities of the classifier when trained using variable combinations suggested by less expensive methods such as sequential forward selection, chi-squared and K-S testing, and physicist intuition. Even in this age of grid computing, the total number of variables which can be tested is limited. In our case, we were limited to considering eleven variables. A less expensive method of variable selection would not only free up computing resources, it would enable us to consider a much larger set of variables for use in the multivariate classifier.

Dennis Mackin
Rice University